CLAIMS

What is claimed is:

1.   A concept classification system that classifies concepts for a domain according to the statistical significance of these concepts in the domain comprising:

at least one general semantic lexicon, implemented as a computational database and accessible by the software modules of the classification system;

at least one text corpus, containing texts specific for a domain for an application of a telephone conversation system, implemented as a computational database; and

a statistical inference algorithm, which makes use of distributional patterns of words in the corpus, together with their conceptual categories in the general semantic lexicon, to derive significant concepts of a particular domain.

2.   A fuzzy grammar rule derivation system to learn fuzzy grammar from corpora comprising:

a shallow parsing algorithm, which assigns syntactic categories to words and segments a sentence into syntactic phrases;

a syntactic weight assignment algorithm, which assigns syntactic weight to context vectors according to their relative importance to a concept; and

a statistical parameter calculation algorithm, which assigns frequency and differentiation parameters to context vectors of concepts.

3.   A concept derivation system, which applies fuzzy grammar rules to a sentence to derive the intended meanings of the speaker comprising:

a matching algorithm to match words to possible multiple concepts, as described in the fuzzy grammar;

a fuzzy inference algorithm which applies fuzzy grammar rules to the context vector of a possible concept and assigns a membership degree to the concept;

an evaluation algorithm, which examines the test results of the fuzzy grammar rule engine and the inference engine to decide if further training of the system is needed or the system has passed the requirement for delivery as the natural language understanding component of a telephone conversation system.